Appendix R

Preliminary Categorization of Tier 1 Screens and Tier 2 Tests by the Screening and Testing Work Group

As mentioned in Chapter Five, the screens and tests being recommended by the EDSTAC vary considerably in terms of the effort necessary to be fully validated and standardized. During their deliberations, the Screening and Testing Work Group (STWG) attempted to categorize the levels of validation required of the recommended screens and tests and their preliminary efforts are found in this Appendix. This information is included to give the reader a sense of where the recommended screens and tests may exist across the spectrum of validation; however, the EDSTAC agreed not to try to reach consensus agreement on where all the screens and tests lie, as they will all need to validated and standardized before being included in the Endocrine Disruptors Screening and Testing Program (EDSTP). The EDSTAC recommends that EPA update this categorization scheme as part of their validation and standardization program.

The recommended screens and tests (including all endpoints) will have to meet all the criteria of relevance and reliability for use in regulatory toxicity screening or testing for Estrogen, Androgen, and Thyroid (EAT) in order to be considered fully validated and standardized (ICCVAM, 1996; Zeiger, 1998). As screens and tests become fully validated and standardized, they will warrant inclusion in the EDSTP according to their specific and appropriate use. None of the screens, new tests, or enhancements to existing test guidelines included in Tier 1 Screening (T1S) or Tier 2 Testing (T2T) completely fulfill these criteria to date. As mentioned throughout Chapter Five, each assay and test under consideration in T1S or T2T needs some level of standardization, validation, methods development, or further research before being accepted as a regulatory toxicity screen or test. The level of standardization and validation varies according to a variety of criteria applied to each of the assays, including: period of time in use, existing level of general acceptance in the endocrine toxicology field, and existing understanding of relevancy and reliability.

The STWG placed the proposed screens and tests in one of four general categories with regard to the level of validation, standardization, or methods development required. A fifth category, discussed in Chapter Five, also identified assays requiring further research.

Category I:

Screens and tests which have been fully validated and standardized are placed in Category I. These procedures meet all the criteria of relevance and reliability for use in regulatory toxicity screening or testing for estrogen, androgen, and thyroid. As other procedures become sufficiently standardized and validated to warrant inclusion in Category I, such screens and tests should be incorporated into the EDSTP according to their specific and appropriate use. Only the following tests are included in Category I:

- Two-Generation Mammalian Reproductive Toxicity Study (1996 Public Draft Guidelines and 1997 TSCA Final Guidelines)
- One-Generation Test

Category II:

Screens and tests which have been in use for a sufficient period of time and which have gained sufficient general acceptance within the field of endocrine toxicology to be considered *de facto* validated (reliable *and* relevant) are included in Category II. These assays measure relevant endpoints, are responsive to endocrine active compounds with a high degree of specificity, are sufficiently sensitive to identify all known active agents, and can reasonably be expected to give reproducible results from laboratory to laboratory, assuming a general level of competence and expertise. Nonetheless, variations in protocols for these screens and tests can produce disparate results. Therefore, standardization of the protocol to be recommended for these screens and tests should be accomplished by EPA before these assays are implemented as screening requirements for endocrine activity or disruption. The following screens and tests are included in Category II:

- ER Binding Assay
- AR Binding Assay
- Rodent 3-Day Uterotrophic Assay (Subcutaneous)
- Rodent 5-7 Day Hershberger Assay
- Rodent 3-Day Uterotrophic Assay (Intraperitoneal);
- Avian Reproduction Test (with Bobwhite Quail and Mallard) (as currently performed)
- Fish Life Cycle Test (Fathead Minnow) Test (as currently performed)
- Mysid Life Cycle Test (Americamysis)

Category III:

Screens and tests which have sufficiently broad use to be generally considered relevant OR reliable to either screening for endocrine activity (Tier 1) or to testing for adverse endocrine-mediated effects (Tier 2) are included in Category III. These assays cannot, however, be generally considered to be both relevant *and* reliable. The level of performance that can be expected of these assays with respect to identifying endocrine active agents or endocrine disruptive effects of chemicals must be clarified. Therefore, these assays should undergo further but focused validation and standardization to define their relevance and reliability for the task of endocrine disruptor screening or testing. The validation required may be focused to answer specific questions about relevance and to provide information regarding specificity and sensitivity. The following screens and tests are included in Category III:

- ER Transcriptional Activation Assay
- AR Transcriptional Activation Assay
- Steroidogenesis Assay with Minced Testis
- Rodent 20-Day Pubertal Female Assay With Thyroid

- Placental Aromatase Assay
- Rodent 14-day Intact Adult Male Assay with Thyroid
- Rodent 20-day Thyroid/Pubertal Male Assay
- Alternative Mammalian Reproduction Test
- Avian Reproduction Test - when performed in multiple generations
- Turtle Egg Assay

Category IV:

Screens and tests which may have relevance to the task of either screening for endocrine activity or testing for endocrine disruptive effects, but whose performance in identifying endocrine active agents or endocrine disruptive effects has seen only limiting testing are included in Category IV. Questions as to whether these assays measure endpoints that are relevant to endocrine activity or endocrine disruptive effects, whether these assays respond with specificity and sensitivity to known endocrine active agents, or whether they identify endocrine disruptive effects cannot be addressed with information currently available. In addition, questions regarding the specific protocols and conditions under which the assays should be conducted must be answered before relevance and reliability can be assessed. Nonetheless, these assays would have sufficient utility, if further developed and validated, to enhance or augment the screening and testing program. The following screens and tests are included in Category IV:

- Frog Metamorphosis Assay
- Fish Gonadal Recrudescence Assay
- 14-Day (PND 9-22) Developmental/Thyroid Assay